

TARTU ÜLIKOOL  
Loodus- ja täppisteaduste valdkond  
Keemia instituut

Kertu Hernits

**Ligniini laguproduktide jaotuskoefitsientide eksperimentaalne määramine  
ja modelleerimine COSMO-RS meetodiga**

Bakalaureusetöö (12EAP)

Keemia

Juhendajad:  
Sofja Tšepelevitš, MSc  
Prof. Ivo Leito

Tartu 2017

## Sisukord

Kasutatud lühendid .....	3
1 Sissejuhatus.....	4
2 Kirjanduse ülevaade.....	5
2.1 Ligniin, selle struktuur ja rakendused.....	5
2.2 Jaotuskoefitsiendid ja nende rakendused.....	6
2.3 Jaotuskoefitsientide eksperimentaalse määramise meetodid.....	7
2.4 Jaotuskoefitsientide arvutusliku ennustamise meetodid.....	8
2.5 COSMO-RS meetod.....	9
3 Eksperimentaalne osa .....	11
3.1 Kasutatud kemikaalid ja töövahendid.....	11
3.2 Eksperimendi kirjeldus .....	11
3.3 Võimalikud määramatuse allikad .....	12
3.4 Vahetult mittemõõdetavate $\log D$ väärtuste leidmine .....	13
4 Arvutused.....	15
4.1 Kasutatud tarkvara ja parameetrid.....	15
4.2 Jaotuskoefitsientide arvutus.....	15
5 Tulemused ja arutelu.....	16
5.1 Kogutud andmestiku kirjeldus.....	16
5.2 Saadud andmete usaldusväärsuse hinnang .....	17
5.3 Hinnang faaside tasakaaluliste koostiste ennustamisele.....	19
5.4 Arvutatud $\log D$ väärtuste üldine kokkulangevus eksperimentaalsete andmetega .....	20
5.5 Arvutuslike ja eksperimentaalsete $\log P$ väärtuste kokkulangevus erinevate orgaaniliste lahustite korral.....	21
5.6 Arvutuslike ja eksperimentaalsete $\log P$ väärtuste kokkulangevuse hinnang erinevate soluutide korral .....	24
6 Kokkuvõte.....	25
7 Summary.....	26
Kasutatud kirjandus .....	27
Lisad.....	30
Lisa 1. Eksperimentaalselt leitud $\log D$ väärtused.....	30
Lisa 2. COSMO-RS meetodiga arvutatud $\log P$ väärtused.....	32
Lisa 3. Kasutatud kemikaalide päritolu ja puhtus .....	34
Infoleht.....	35
Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks ....	36

## Kasutatud lühendid

COSMO-RS	–	<i>Conductor-like Screening Model for Realistic Solvation</i>
DFT	–	<i>Density Functional Theory</i> - tihedusfunktsionaali meetod
log <i>D</i>	–	jaotuskoefitsiendi logaritm (arvestades soluudi ionisatsiooni)
log <i>P</i>	–	jaotuskoefitsiendi logaritm (neutraalse soluudi jaoks)
MTBE	–	Metüül- <i>tert</i> -butüüleeter
PCA	–	<i>Principal Component Analysis</i> - peakomponentide analüüs
QSPR	–	<i>Quantitative Structure-Property Relationship</i> - kvantitatiivne struktuuri ja omaduse sõltuvus
RPLC	–	<i>Reversed-Phase Liquid Chromatography</i> - pöördfaas-vedelikkromatograafia

## 1 Sissejuhatus

Ligniin on üks levinumaid biomassi komponente ja põhiline taasutuv aromaadne ressurss looduses [1]. Ühtlasi on ligniin perspektiivne lähteaine kõrge lisandväärtusega kemikaalide ja kõrge kvaliteediga biokütuste tootmiseks, kuna seda tekib tselluloosi tootmisel suures koguses [2]. Enamasti ei kasutata ligniini otstarbekalt [3], kuid sellest oleks võimalik toota väärtuslikke madalmolekulaarseid ühendeid [4]. Ligniini ümbertöötlemiseks ning tekkinud monomeeride segust huvipakkuvate ühendite eraldamiseks kasutatakse muuhulgas vedelik-vedelik ekstraktsiooni [5]. Meetodi tulemuslikkus sõltub soluudi jaotustasakaalust lahustite vahel ning seega on kasulik teada ainete jaotuskoefitsiente erinevates lahustitepaarides.

Jaotuskoefitsiente on võimalik määrata eksperimentaalsete meetoditega [6–8] (näiteks otseselt loksutusmeetodiga või kaudselt pöördfaas-vedelikkromatograafia abil) kui ka ennustada arvutuslikult [9] (näiteks QSPR mudelite või COSMO-RS meetodi abil). Kuna ainete jaotumise eksperimentaalne uurimine on töö-, aja- ja ressursimahukas protsess, on tihti eelistatud usaldusväärse arvutusliku meetodi kasutamine. Teiselt poolt, arvutusmeetodite parametrizeerimiseks ja/või valideerimiseks on vaja usaldusväärseid eksperimentaalseid andmeid.

Käesoleval tööl on kaks peamist eesmärki: (1) luua eksperimentaalselt leitud jaotuskoefitsientide väärtuste täielik andmematriks, kaasates ligniiniist potentsiaalselt saadavaid ja nendega struktuurselt sarnaseid madalmolekulaarseid ühendeid ning erineva struktuuri ja polaarsusega lahusteid, ja (2) võrrelda leitud jaotuskoefitsiente vastavate COSMO-RS arvutusmeetodiga arvutatud väärtustega, et hinnata meetodi ennustusvõimet erinevate solventide ja ainete korral.

Saadud eksperimentaalsed andmed on potentsiaalselt kasutatavad nii otseselt ekstraktsiooni kavandamiseks kui ka erinevate arvutusmeetodite ja mudelite arendamiseks ja valideerimiseks.

## 2 Kirjanduse ülevaade

### 2.1 Ligniin, selle struktuur ja rakendused

Ligniin on keerulise struktuuriga vees lahustumatu biopolümeer. See on biomassi oluline makromolekulaarne komponent. Olenevalt puuliigist moodustab ligniin 15-30% puidu massist. Põhiliselt asetseb ta kiurakkude vahel ja tagab puidu struktuurse jäikuse [5].

Ligniin tekib taimedes biosünteesi käigus monolignoolidest, mille põhilised esindajad on hüdroksü- ja metoksü-asendatud kaneelalkoholi (3-fenüülprop-2-een-3-ooli) molekulid. Erinevalt tselluloosist, kus  $\beta$ -D-glükopüraanoosi jäägid on seotud korrapäraselt  $\beta$ -1,4-glükosiidsidemega, on ligniin ebakorrapärase struktuuriga: see koosneb erinevatest alamstruktuuridest, mis esinevad ja korduvad juhuslikult [10]. Ligniin tekib kõrvalsaadusena tselluloosi tootmisel puidust ning seda ei kasutata või kasutatakse ebaotstarbekalt [3].

Ligniinist võib toota palju väärtuslikke madalmolekulaarseid ühendeid, nt fenooli ja süsivesikuid [4]. Ligniini lagundamiseks mono- ja oligomeerideks kasutatakse katalüüsi ja/või kõrget temperatuuri [10], mille järel eraldatakse saadud produktide segust huvipakkuvad ühendid. Üks näide ligniinist monomeeride tootmise meetodist on alusekatalüütiline depolümerisatsioon [5]. Põhilised monomeerid mida saadakse selle meetodiga on fenool, guajakool, katehhool, 4-metüülkatehhool, 3-metoksükatehhool, 4-etüülkatehhool, süringool, vanilliin, apotsüniin (4-hüdroksü-3-metoksüatsetofenoon), pürgallool, sirelaldehüüd ja atsetosüringoon [5]. Monomeeride eraldamine on keeruline protsess, mida on tööstuslikul tasemel raske täielikult ning efektiivselt teostada. Samas pole see võimatu - maailmas laialdaselt tegutsev ettevõtte Borregaard toodab puidust vanilliini (protsessi detailne kirjeldus pole vabalt kättesaadav) ning kasutab seda näiteks toidu valmistamiseks.

Monomeeride eraldamise protseduur peale depolümerisatsiooni võib olla erinevate meetodite kombinatsioon. Nendeks võivad olla näiteks vedelik-vedelik ekstraktsioon, vaakumdestillatsioon, kromatograafia ja kristallisatsioon [5]. Vedelik-vedelik ekstraktsioon on kasulik meetod kõrge viskoossusega, kõrgema molekulmassiga ja termiliselt ebapüsivate ühendite (sh oligomeeride) eemaldamiseks segudest. Termiliselt ebapüsivate ühendite eemaldamiseks segudest saab kasutada ka kiirkromatograafiat, kuid suuremahulise tööstuse korral on meetod liiga kallis [5]. Vedelik-vedelik ekstraktsiooni tulemuslikkus sõltub soluudi jaotuskoefitsiendist solventide vahel. Mida suurema väärtusega on jaotuskoefitsient, seda

efektiivsem on ekstraheerimine. Seega on ainete eraldamise kavandamiseks kasulik teada nende jaotuskoefitsiente erinevates solvendipaarides.

## 2.2 Jaotuskoefitsiendid ja nende rakendused

Aine jaotumist kahe mitteseguneva faasi vahel kirjendatakse jaotuskoefitsiendiga. See näitab soluudi tasakaaluliste kontsentratsioonide suhet kahes faasis, mis on põhjustatud ühendi erinevast lahustuvusest neis. Jaotuskoefitsiente on kahte tüüpi.  $P$  (*partition coefficient*) näitab ioniseerimata soluudi vormide tasakaaluliste kontsentratsioonide suhet (võrrand 1).  $D$  (*distribution coefficient*) arvestab nii ioniseeritud kui ioniseerimata vormide jaotust (võrrand 2) [6].

$$P = \frac{[\text{neu}]_{\text{org}}}{[\text{neu}]_{\text{vesi}}} \quad (1)$$

$$D = \frac{[\text{neu}]_{\text{org}} + [\text{ioon}]_{\text{org}}}{[\text{neu}]_{\text{vesi}} + [\text{ioon}]_{\text{vesi}}} \quad (2)$$

Kus

$P$  – ioniseerumata soluudi tasakaaluliste kontsentratsioonide suhe faaside vahel;

$[\text{neu}]_{\text{org}}$  – ioniseerumata soluudi tasakaaluline kontsentratsioon orgaanilises faasis;

$[\text{neu}]_{\text{vesi}}$  – ioniseerumata soluudi tasakaaluline kontsentratsioon veefaasis;

$D$  – soluudi ioniseerunud ja ioniseerimata vormide kontsentratsioonide summa suhe faaside vahel;

$[\text{ioon}]_{\text{org}}$  – ioniseerunud soluudi kontsentratsioon orgaanilises faasis;

$[\text{ioon}]_{\text{vesi}}$  – ioniseerunud soluudi kontsentratsioon veefaasis.

Aastakümneid on uuritud looduslike ühendite, näiteks taimeekstraktide komponentide, jaotumist erinevate lahustite vahel. Tänapäeval kasutatakse jaotuskoefitsiente molekulivaheliste jõudude mõistmiseks. Jaotuskoefitsiendid annavad infot interaktsioonide kohta molekulide vahel [6].

Praktilisest küljest on jaotuskoefitsientidel väga suur tähtsus farmakoloogias. Ravimi lipofiilsus, mida väljendatakse enamasti oktanool-vesi jaotuskoefitsiendi kaudu, mõjutab tugevalt ravimi liikumist ja muundumist organismis ja seeläbi ka seda, kui kergesti ravim täidab oma eesmärgi, kui tugevat efekti põhjustab ja kui kaua püsib kehas aktiivsena. Seega on molekuli  $\log P$  oluliseks parameetrigi, mida jälgitakse uute ravimite arenduses hindamaks uuritavate ühendite sobivust ravimikandidaatideks [11,12].

Jaotuskoefitsiente kasutatakse ka keskkonnauuringutes. Aine jaotumist kahefaasilises süsteemis vaadeldakse hüdroboosuse hindamiseks. Seda kasutatakse näiteks radionukliidide liikuvuse uurimiseks põhjavees ja bioakumulatsiooni ennustamiseks. Hüdrogeoloogias kasutatakse jaotuskoefitsiente hüdrofoobsete ühendite migratsiooni ennustamiseks pinnases ja põhjavees [13].

### **2.3 Jaotuskoefitsientide eksperimentaalse määramise meetodid**

Põhiline meetod jaotuskoefitsientide eksperimentaalseks määramiseks on nn loksutamismeetod (*shake-flask method*) [7]. See meetod põhineb otseselt vedelik-vedelik ekstraktsioonil. Soluudi lahust ühes lahustis loksutatakse koos teise, esimesega mitteseguneva, lahustiga, mille käigus toimub soluudi jaotumine kahe faasi vahel. Sellele järgneb lahuste eraldamine ning soluudi kontsentratsiooni määramine ühes või mõlemas faasis [6].

Jaotuskoefitsiendi ligikaudseks hindamiseks võib kasutada soluudi lahustuvuste suhet kahes kasutatavas solvendis. Peab aga arvestama, et sel moel leitud  $P$  väärtus on tegelikkusele lähedane vaid siis, kui nii soluudi lahustuvus solventides kui ka solventide lahustuvus teineteises on piisavalt madalad, et küllastatud lahuste omadused ei erineks oluliselt puhaste lahustite omadustest. See pole alati nii. Vee lahustuvus paljudes lahustites on üsna kõrge ning vee lisand muudab arvestatavalt solventide iseloomu. Teisalt, paljude ainete lahustuvus orgaanilistes solventides on kõrge ja soluudi olemasolu muudab ka orgaanilise faasi iseloomu/omadusi [6].

Otsese eksperimentaalse jaotuskoefitsientide määramisega võivad kaasneda mitmed probleemid. Pindaktiivsete omadustega soluudid võivad põhjustada emulsiooni tekkimise, mille tagajärjel vedeliku kihid ei eraldu. Emulsiooni lõhkumiseks kasutatakse tsentrifuugimist, pikaajalist seista laskmist või nende võtete kombinatsiooni. Lisaks on võimalik kasutada difusioonil põhinevat meetodit. Paraku on see ajakulukas ning enamasti kasutatakse meetodit koos teiste võtetega [6].

On olemas mitu tehnilist lahendust kahe faasi tasakaalustamiseks ja eraldamiseks. Kahe faasi vaheline tasakaal püstitub juba 1-5 minutiga, kui viia mõlemad faasid suletud katseklaasi ja seda 50-100 korda ümber pöörata. Teised segamismeetodid võivad nõuda rohkem aega – mõnikümmend [14] kuni mõnisada [15] minutit. Tuleb vältida väga tugevat loksutamist, kuna see võib viia emulsiooni tekkeni. Emulsiooni ei saa alati tuvastada visuaalselt (mõlemad

faasid võivad välja näha selged), seega usaldusväärsete tulemuste saamise jaoks on soovitatav alati läbi viia tsentrifuugimine [6].

Väga madala kontsentratsiooniga ( $<10^{-6}$  M) lahuste korral võib probleemiks olla jaotuvate ainete adsorptsioon anumate seintele [6].

Jaotuskoefitsiente on võimalik määrata ka kaudselt pöördfaas-vedelikkromatograafia (*RPLC*, *Reversed-Phase Liquid Chromatography*) abil. Vedelikkromatograafia on kiire ja lihtne meetod, mis kasutab ainete eraldamiseks statsionaarset ja mobiilset faasi. Statsionaarseks faasiks ehk sorbendiks ehk täidiseks on kolonnis paigal seisev poorne materjal ning mobiilseks faasiks on vedelik, mis voolab läbi poorse materjali ning kannab uuritavaid ained edasi. Pöördfaas-vedelikkromatograafia korral on statsionaarse faasi pinnale immobiliseeritud vähepolaarsed hüdrofoobsed rühmad (enamasti pikad alifaatsed ahelad) ja mobiilseks faasiks on suhteliselt polaarne lahusti (nt vee ja metanooli segu). Ainete lahutumine sõltub sellest, kui tugevalt nad mobiilse ja statsionaarse faasiga interakteeruvad. Sellest tulenevalt liiguvad ained kolonnis erineva kiirusega [8].

Jaotuskoefitsientide kaudsel määramisel eeldatakse, et soluudi jaotuskoefitsiendi logaritm on lineaarses sõltuvuses tema retentsioonifaktori logaritmist (võrrand 3):

$$\text{Log}D = a \cdot \log k + b \quad (3)$$

Kus  $a$  ja  $b$  on regressioonikoefitsiendid,  $k$  on retentsioonifaktor ja  $D$  on jaotuskoefitsient. Koefitsiendid  $a$  ja  $b$  leitakse, mõõtes sarnaste soluutide (mille  $\log D$  väärtused on teada) retentsioonifaktorid RPLC kolonnis. Korraga saab mõõta mitut ühendit ning soluudi kontsentratsiooni pole vaja eelnevalt teada. On oluline, et mobiilne faas sisaldaks vähemalt 25% vett, korrelatsiooni moodustaksid vähemalt 6 struktuurselt sarnast ainet ja kalibratsioonivahemikust väljapoole võib ekstrapoleerida vaid äärmisel vajadusel (näiteks ülikõrgete  $\log D$  väärtustega ainete korral) [7]. Meetodi miinuseks on see, et leitud  $\log D$  väärtus sõltub korrelatsioonikoefitsientide määramise usaldusväärsusest [7].

## 2.4 Jaotuskoefitsientide arvutusliku ennustamise meetodid

Jaotustasakaalude arvutuslik ennustamine on huvipakkuv väga erinevate uuringute seisukohalt. See on kiirem ja keskkonnasõbralikum kui eksperimentaalne määramine ja lisaks võimaldab hinnata seni sünteesimata või puhtal kujul eraldamata soluutide omadusi.



Üks laialt levinud meetod  $\log P/\log D$  väärtuste ennustamiseks on QSPR mudelid. Kvantitatiivne struktuuri ja aktiivsuse/omaduse sõltuvus (QSPR) kasutab multilineaarse regressiooni meetodit, et analüüsida suures andmekogumis erinevaid molekulaarseid struktuure ja siduda aine füsiokeemilisi omadusi või bioloogilist aktiivsust aine keemilise struktuuriga. QSPR mudeli abil saab ennustada ka eksperimentaalselt mitte kättesaadavate ühendite omadusi [9].

Mudeli loomiseks kogutakse andmeid ainete bioloogiliste omaduste kohta (teostatakse mõõtmised või kogutakse andmeid kirjandusest). Seejärel leitakse igale ainele sobivad deskriptorid ehk struktuuri kirjeldavad parameetrid. Nendele sammudele järgneb omaduste modelleerimine, milleks kasutatakse statistilisi meetodeid. Mudeli loomisel on väga oluline valideerimine, millega hinnatakse selle ennustusvõimet. On oluline, et mudel oleks statistiliselt korrektne ning kasutatav oma defineeritud rakendusala piires. Võimalik on rakendada ka väga erinevatele ainetele, kuid seda on raske saavutada [9].

QSPR mudelitel on ka olulised puudused. Nimelt ei saa uurida ainete käitumist uudsetes solventides, mille kohta pole eksperimendiandmeid, suvalistes solvendisegudes ning ainulaadsete ainete (millele sarnased molekulid ei ole mudeli loomiseks kasutatud andmestikus piisaval määral esindatud) omaduste ennustamine pole suure tõenäosusega täpne [16,17].

## 2.5 COSMO-RS meetod

COSMO-RS (*Conductor-like Screening Model for Realistic Solvation*) [18,19] on meetod ainete termodünaamiliste omaduste ennustamiseks vedelfaasis. Meetod põhineb kvantkeemia ja statistilise termodünaamika kombinatsioonil, mis võimaldab tõhusalt arvutada omadusi, mida vähesed teised meetodid suudavad [20].

COSMO-RS rakendamise esimese sammuna arvutatakse tihedusfunktsionaali (DFT) meetodil molekulide geomeetriad, koguenergiad lõpmata suure dielektrilise konstandiga keskkonnas ning leitakse elektrostaatilise potentsiaali jaotus molekuli pinnal (sigma-potentsiaali jaotus). Molekulidevaheliste vastasmõjude energia vedelas segus leitakse pinnasegmentide vaheliste elektrostaatiliste vastasmõjude kaudu. Lõpptulemusena saadakse segus olevate komponentide keemilised potentsiaalid ning seeläbi arvutatakse termodünaamilisi parameetreid. COSMO-RS ei vaatle terveid molekule, vaid pinnasegmente eraldi ning sellest tulenevalt võivad mõningad geomeetrilised ja steerilised efektid kaduda [20]. Näiteks võib tekkida olukord, kus

COSMO-RS meetod arvestab teatud pinnasegmentide vastasmõjusid keskkonnaga, kuid tegelikkuses pole teiste molekulide ligipääs neile pinnasegmentidele võimalik asendusrühmade paiknemise tõttu (steerilised takistused).

COSMO-RS meetodis on ületatud paljud teistele ennustusmeetoditele omased puudused. Esiteks, see meetod ei tee vahet lahustil ja lahustunud ainel ja seeläbi võimaldab töötada lisaks lahjadele lahustele (millega on piiratud enamik muid vedelfaasis töötavaid arvutusmeetodeid) ka mitmekomponendiliste segudega, kus komponentide sisaldused on võrreldavad. Teiseks, on võimalik ennustada mistahes molekulstruktuuriga solventide omadusi. Saadud tulemused ei lange alati kokku vastavate eksperimentaalsete andmetega, kuid ennustused suhtelisel skaalal on enamasti korrektsed ning annavad palju kasulikku infot [20].

Üheks meetodi piiranguks on suutmatus korrektselt esitada sekundaarsete ja tertsiaarsete amiinide vastasmõjusid vesiniksidemeid andvate soluutide ja solventidega. Teine oluline piirang esineb keemilise potentsiaali täpsel arvutamisel väikeste molekulide korral. See välistab COSMO-RS kasutamist väga suurt täpsust nõudvates rakendustes. Kõige häirivamaks probleemiks on piirangud ioonidega (eriti anioonidega) töötamisel. Selle põhjuseks on nn „väljaulatuv laeng“, mille korral mingi osa aniooni elektronkattest on jäetud arvestamata kvantkeemilise arvutuse etapil. Sellest tulenevalt on anioonide omaduste ennustamisel tulemuste kvaliteet madalam kui neutraalsete või katioonsete osakeste korral [20].

COSMO-RS meetodit kasutatakse nii ravimite väljatöötamisel, põllumajanduses kemikaalijääkide saastuse uurimisel kui ka aur-vedelik diagrammide koostamisel ning erinevate termodünaamiliste parameetrite (näiteks jaotuskoefitsientide) arvutamiseks [20–22]. Paljud maailmas tuntud ettevõtted (näiteks Bayer, Pfizer) kasutavad COSMOLogic tehnoloogiaid oma igapäevases töös.

COSMO-RS täpsuse hinnanguks jaotuskoefitsientide arvutamisel on 0,35-0,40 log ühikut (parametriseerimises kasutatud eksperimentaalsete andmete ja vastavate ennustuste ruutkeskmise hälve) [23].

### 3 Eksperimentaalne osa

#### 3.1 Kasutatud kemikaalid ja töövahendid

Soluutide valikul sadadest teadaolevatest ligniini laguproduktidest ja neile struktuurselt sarnastest ainetest lähtuti järgmistest praktilistest kaalutlustest: aine kättesaadavus sobiva puhtuse ja hinnaga, UV-Vis aktiivsus (võimaldamaks kasutada UV-Vis detektorit), stabiilsus lahustes,  $\log D$  väärtused kasutatud metoodika rakendusala piires (vahemikus ligikaudu -3 kuni 3). Lisaks peeti silmas seda, et uuritavad ained oleksid omaduste poolest võimalikult mitmekesised.

Töös uuritud ained olid furfurool, fenool, 2-isopropüülfenool, 3-metüülfenool, katehhool, 4-metüülkatehhool, resortsinool, 1,3-dimetoksübenseen, guajakool, vanilliin, isovanilliin, koniferüülaldehüüd, hüdrokinoon, 1,4-dimetoksübenseen, 4-aniisalaldehüüd, 1,2-dimetüülbenseen, atsetofenoon, 2'-hüdroksüatsetofenoon, antrakinoon, 9-hüdroksüfluoreen, pürogallool, 3-metoksükatehhool, süringool, sirelalaldehüüd, sirelhape, gallushape, vanilliinhape, 4-hüdroksübensoehape, trans-feruulhape ja 4-aniishape.

Solventide valikul võeti arvesse nende omaduste mitmekesisust (erinev polaarsus, struktuur), kättesaadavust sobiva puhtuse ja hinnaga, toksilisust ja lenduvust.

Töös kasutatud solvendid olid *n*-heksaan, toluen, klorobenseen, diklorometaan, metüül-*tert*-butüüleeter, isopropüülatsetaat, bensüülalkohol, tsükloheksanool, *n*-butanool ja *n*-oktanool.

Kasutatud kemikaalide päritolu ja puhtus on toodud Lisas 3.

#### 3.2 Eksperimendi kirjeldus

Vedelik-vedelik ekstraktsiooni läbiviimiseks kaaluti 1-2 mg (vedelike korral 1-2  $\mu$ l) analüüti ning lahustati 10 ml-s Milli-Q vees või 0,1 M HCl lahuses. Ainete kiiremaks lahustumiseks kasutati vajadusel ultrahelivanni. Seejärel pipeteeriti standardsetesse 2 ml krimp-korgiga viaalidesse võrdsed kogused uuritavaid aineid sisaldavat vesilahust ning orgaanilist lahustit. Pipeteeritud kogused jäid vahemikku 0,6-0,7 ml, kuid enim kasutatud vedelike ruumala oli 0,6 ml. Tõstmaks töö efektiivsust lahustati ühes segus enamasti 5-7 uuritavat ainet. Happelisi soluute sisaldavad neutraalsed lahused puhverdati. pH reguleeriti väärtusele  $7,0 \pm 0,1$  Na<sub>2</sub>HPO<sub>4</sub> abil.

Seejärel viaalid suleti, asetati hoidjasse ja kinnitati loksutile. Loksutamist viidi läbi 1,5 tundi kiirusega umbes 250 ringi minutis. Minimaalne vajalik loksutuaeg oli meie uurimisgrupis juba varem välja selgitatud [24].

Viimase etapina paigutati viaalid vedelikkromatograafi automaatproovivõtjasse ning teostati nii orgaanilise kui ka veefaasi kromatograafilise analüüs. Selleks kasutati vedelikkromatograafi firmalt Agilent ja kolonni Agilent Eclipse XDB-C18 mõõtmetega 4,6 x 250 mm ning 5 µm täidiseosakese suurusega. Detektorina kasutati viiekanalist UV-Vis detektorit. Aparaat seadistati sobivate parameetritega (süsti ruumala, nõela liikumine vedeliku kihtides viaalis, detekteerimise lainepikkused). Süsti ruumalaks valiti orgaanilise faasi korral 5 ja veefaasi korral 30 µL, kuna see andis sobivad piigi suurused hilisemaks analüüsiks. Detekteerimiseks kasutati lainepikkusi 220, 245, 254, 274, ja 280 nm, kuna praktiliselt kõik uuritud solutide olulised neeldumismaksimumid paiknevad selles lainepikkuste vahemikus. Saadud piikide pindalade järgi arvutati uuritavatele ainete  $\log D$  väärtused (võrrand 4).

$$\log D = \log \left( \frac{V_{\text{vesi}} \times S_{\text{org}}}{V_{\text{org}} \times S_{\text{vesi}}} \right) \quad (4)$$

Kus

$V_{\text{vesi}}$  – süsti ruumala veefaasi analüüsil;

$S_{\text{org}}$  – analüüdi piigi pindala orgaanilise faasi kromatogrammil;

$V_{\text{org}}$  – süsti ruumala orgaanilise faasi analüüsil;

$S_{\text{vesi}}$  – analüüdi piigi pindala veefaasi kromatogrammil.

Käesolevas töös kasutati lisaks autori kogutud andmetele ka teiste uurimisgrupi liikmete (Sofja Tšepelevitš, Jaroslav Jenčo) kogutud andmeid, mida oli kokku umbes 30% kogu andmestikust.

### 3.3 Võimalikud määramatuse allikad

Oluline on tähelepanu pöörata võimalikele määramatuse allikatele. Üheks probleemiks eksperimendi teostamisel võib olla temperatuuri kõikumine. Kogu eksperimendi käigus jälgiti laboriruumi temperatuuri, sest ainete jaotumine faaside vahel sõltub oluliselt temperatuurist. Ruumi temperatuur jäi 20°C ja 24°C vahele, temperatuur automaatproovivõtjas, kus viaalid seisisid kuni 13 tundi enne analüüsi, oli ruumitemperatuurist ligikaudu 3 kraadi kõrgem.

Kirjandusest on teada, et temperatuuri efekt  $\log P/\log D$  väärtustele on suurusjärgus 0,01 log ühikut  $1^{\circ}\text{C}$  kohta (täpne väärtus sõltub soluudist ja solventidest) [25]. See on suhteliselt väike muutus, seega käesolevas töös ei rakendatud täiendavat temperatuuri kontrolli (termostateerimist).

pH muutuste vältimiseks puhverdati happeid sisaldavad veefaasid. Soluutide ionisatsioon võib muuta lahuse pH väärtust, mis omakorda võib mõjutada analüütide jaotustasakaale. Kui lahuse pH muutub madalamaks, happeliste ühendite  $\log D$  väärtused muutuvad kõrgemaks, ja vastupidi.

Kuna ühes segus sisaldus 5-7 uuritavat ainet, tuleb tähelepanu pöörata ka võimalikele ainetevahelistele vastasmõjudele. Soluutide omavaheliste reaktsioonide käigus võisid moodustuda uued ühendid, mida järelitati tundmatute piikide ilmumisest kromatogrammidele. Sellistes katsetes saadud andmeid kasutati ettevaatusega. Arvesse tuleb võtta ka selliseid tasakaalu mõjutavaid vastasmõjusid, mis ei ole keemiliste reaktsioonide põhjustatud – nt tugevate vesiniksidemete teket soluutide vahel. Näiteks on teada, et karboksüülhapped moodustavad vähepolaarsetes lahustes dimeere [6]. Kuna dimeeri jaotuskoefitsient on oodatavasti neutraali jaotuskoefitsiendist kõrgem, mõjutab selline lisavastasmõju mõõdetavat  $\log P/\log D$  väärtust. Ka käesolevas töös esines olukordi, kus  $\log D$  väärtused erinesid oluliselt eelnevates katsetes saadud tulemustest. Võimalikuks põhjuseks on erinevad soluutide kombinatsioonid neis katsetes ja järelikult ka erinevad soluutide vahelised vastasmõjud. Selle mõju efektist vabanemiseks püüti korduskatsetes soluutide kombinatsioone varieerida ning saadud tulemused keskmistati. Lisaks püüti kasutada võimalikult madalaid soluutide kontsentratsioone.

Lisaks on oluline viaalidesse pipeteeritud vedelike ruumala. Efektiivseks segunemiseks peab ruumala olema piisav, kuid ei tohi olla liiga suur. Viaalides peab olema vaba ruumi, et segunemine saaks efektiivselt toimida.

### **3.4 Vahetult mittemõõdetavate $\log D$ väärtuste leidmine**

Eksperimentaalse töö käigus selgus, et kõiki huvipakkuvaid  $\log D$  väärtusi pole võimalik kasutatud meetodika abil eksperimentaalselt määrata. Peamisteks põhjusteks on soovimatud reaktsioonid lahustes (eeldatavasti uuritavate soluutide ja solvendi lisandite vahel) või kordusmõõtmiste tulemuste suur lahknevus, mis võib viidata tuvastamata probleemidele. Lisaks on osade ainete  $\log D$  väärtused mõnedes lahustipaarides väljaspool meetodika

rakendusala ( $\log D$  vahemikku, milles metoodika annab usaldusväärseid tulemusi).

Kuigi puuduvad eksperimentaalsed väärtused ei takista otseselt andmete võrdlust COSMORS tulemustega, oli töö eesmärgiks täieliku andmematriksi saamine ja selleks leiti puuduvad andmed statistilise meetoditega. Täielik, sama metoodikaga ja sama uurimisrühma poolt kogutud kooskõlaline andmestik on väga kasulik ennustavate mudelite arendamiseks ja valideerimiseks. Puuduvad väärtused andmematriksis piiravad mõningate statistiliste meetodite rakendatavust. Näiteks peakomponentide analüüsi (PCA) jaoks on väga oluline, et andmematriks oleks täielik, kuna puuduvate väärtuste korral ei ole meetod rakendatav, aga puuduvate väärtuste asendamisel ebakvaliteetsete hinnangutega võib meetod anda ebausaldusväärseid tulemusi.

Enamikul juhtudel ennustati puuduvad  $\log D$  väärtused katseandmetest kasutades lineaarseid korrelatsioone sama ühendi  $\log D$  väärtuste vahel erinevates lahustipaarides või korrelatsiooni erinevate (kuid sarnaste omadustega) ühendite  $\log D$  väärtuste vahel samas lahustipaaris. Juhtudel, kui hapete otseselt mõõdetud  $\log D$  väärtuste reprodutseeritavus oli halb, arvutati neid võrrandi 5 järgi kasutades vastavaid  $\log P$  ja  $pK_a$  väärtusi.

$$\log D = \log P - \log(1 + 10^{\text{pH} - \text{p}K_a}) \quad (5)$$

Mõnede  $\log D$  väärtuste täiustamiseks kasutati peakomponentide analüüsi (PCA, *Principal Component Analysis*). Meetodi käigus luuakse esialgsetest muutujatest uusi – nn peakomponente – mis kujutavad endast algsete muutujate lineaarseid kombinatsioone. Andmestikku on võimalik peakomponentide alusel taastada. Kui kasutada selleks piiratud arvu peakomponente, läheb osa infot (ennekõike juhuslik müra) kaotsi. Seda meetodi omadust kasutati ära, asendades mõned vähese usaldusväärsusega  $\log P/\log D$  väärtused kahe peakomponendi baasil taastatud väärtustega.

## 4 Arvutused

### 4.1 Kasutatud tarkvara ja parameetrid

DFT (*Density Functional Theory*) arvutused uuritavate ainete tasakaalulise geomeetria ja pinnalaengute jaotuse leidmiseks olid meie uurimisgrupis varasemalt tehtud (arvutusmetoodika analoogne allikas [26] kirjeldatule). Jaotuskoefitsientide arvutamiseks statistilise termodünaamika meetodil kasutati COSMO*thermX* tarkvara (Version C3.0 Release 17.01) [27,28] parametrisatsiooniga BP\_TZVPD\_FINE\_C30\_1701. Molekulide kõigi konformeeride suhtelisi kaalusid arvestati vastavalt COSMO-RS tavapärasele lähenemisele. Arvutustes kasutati temperatuuri 23 °C (kui pole teisiti täpsustatud), mis vastab keskmistatud temperatuurile eksperimentide läbiviimise ajal.

### 4.2 Jaotuskoefitsientide arvutus

Jaotuskoefitsiendi arvutus COSMO*thermX* tarkvara abil koosnes kahest etapist. Esimese etapina arvutati kahe lahusti koostised peale vastastikku küllastumist ehk faaside koostised (moolimurdudes), milleks kasutati funktsiooni *Liquid Extraction*. Vee lahustuvus mitmes kasutatud orgaanilises lahustis on küllalt kõrge, lisaks madala polaarsusega lahusti puhul isegi väike veesisaldus võib oluliselt muuta selle omadusi. Seega lahustite vastastikuse küllastumise arvestamata jätmine või ebatäpne ennustamine põhjustab suure tõenäosusega vigu  $\log P$  väärtustes. Esimese etapi täpsust ning ennustusvõimet analüüsiti käesoleva töö alajoatuses 5.3. Teise etapina leiti funktsiooni  $\log P / \log D$  abil soluutide neutraalsete vormide jaotustasakaalu konstantide logaritme ( $\log P$ ), kasutades eelnevas etapis leitud faaside koostisi. Arvutuste teostamisel seati analüütide kontsentratsioonid võrdseks nulliga, imiteerides seega lahjade lahuste kasutamist.  $\log D$  väärtused leiti võrrandi 5 kaudu kasutades eelnevalt arvutatud  $\log P$  ja kirjandusest kogutud  $pK_a$  väärtusi.

## 5 Tulemused ja arutelu

### 5.1 Kogutud andmestiku kirjeldus

Eksperimentaalselt leitud  $\log D$  väärtused on toodud Lisas 1. COSMO-RS meetodiga arvutatud  $\log D$  väärtused on esitatud Lisas 2. Käesolevas töös leiti kokku 600  $\log D$  väärtust 30 erineva uuritava aine jaoks 20 solvendipaaris. Uuritavateks aineteks olid erinevad ligniinist potentsiaalselt saadavad või ligniiniga struktuurselt lähedased ained. Solvendipaaride orgaanilisteks komponentideks olid 10 orgaanilist lahustit (Tabel 1) ja veefaasi-komponentideks olid vesi ja hapestatud (0,1 M HCl) vesi. Orgaaniliste lahustitena kasutati erinevate aineklasside ühendeid, püüdes kaasata erineva polaarsuse ja erinevate struktuursete omadustega lahusteid: nii aroomaatseid kui ka alifaatseid, aprotoonseid ja protoonseid (vesiniksideme doonoreid). Oluline on märkida, et lahustite  $E_T(30)$  väärtused (Tabel 1) vastavad kuivadele lahustitele, kuid ekstraheerimisel kasutatud lahusti oli paratamatult küllastatud veega.

Saadud andmete kvaliteeti mõjutab asjaolu, et kõiki  $\log P/\log D$  väärtusi ei suudetud otseselt mõõta. 15% väärtustest leiti statistiliste meetoditega: 34 väärtust leiti kasutades lineaarseid korrelatsioone sama ühendi  $\log P$  ja  $\log D$  väärtuste vahel; 12 väärtust – kasutades korrelatsioone erinevate ühendite  $\log P/\log D$  väärtuste vahel kahes sobivas solvendipaaris; 8 väärtuse korrigeerimiseks kasutati peakomponentide analüüsi ning osa hapete  $\log D$  väärtustest (6% kogu andmestikust) arvutati võrrandi 5 järgi.

**Tabel 1. Kasutatud orgaaniliste lahustite omadused. Tähistused:  $E_T(30)$  – keskkonna polaarsust iseloomustav parameeter (kcal/mol); HB – vesiniksideme andmise võime (A - aktseptor, D - doonor)**

Lahusti	$E_T(30)$	HB	Lahusti	$E_T(30)$	HB
<i>n</i> -Heksaan	31,0 <sup>a</sup>	-	Isopropüülatsetaat	36-38 <sup>b</sup>	A
Tolueen	33,9 <sup>a</sup>	-	Bensüülalkohol	50,8 <sup>c</sup>	A/D
Klorobenseen	36,8 <sup>a</sup>	-	Tsükloheksanool	46,9 <sup>c</sup>	A/D
Diklorometaan	40,7 <sup>a</sup>	-	<i>n</i> -Butanool	49,7 <sup>d</sup>	A/D
Metüül- <i>tert</i> -butüüleeter (MTBE)	34,7 <sup>e</sup>	A	<i>n</i> -Oktanool	48,3 <sup>c</sup>	A/D

<sup>a</sup> Viide [29]; <sup>b</sup> Eksperimentaalne väärtus puudub kirjanduses, vahemik on hinnatud erinevate äädikhape estrite kirjanduslike [29] väärtuste põhjal. <sup>c</sup> Viide [30]; <sup>d</sup> Viide [31]; <sup>e</sup> Viide [32].



## 5.2 Saadud andmete usaldusväärsuse hinnang

Andmete usaldusväärsuse hindamiseks arvutati eksperimendi päevadevahelist korratavust (juhusliku vea hinnang) ja võrreldi saadud  $\log P$  väärtusi võimaluse korral kirjanduses avaldatud väärtustega (süsteemaatilise vea hinnang). Eksperimendi päevadevaheline korratavus, väljendatuna kordusmõõtmiste tulemuste kogutud standardhälvena, on 0,11 log ühikut (leitud ühtekokku 1664 väärtusest, vabadusastmete arv 1128), mida võib lugeda saadud väärtuste keskmiseks standardmääramatuse hinnanguks. Samas varieerub korratavus solventide kaupa ning tulemused on mõnevõrra paremad protoonsete lahustite korral (Tabel 4). Kirjandusest on enamikule soovitud ainetele võimalik saada vaid oktanool-vesi jaotuskoefitsiente ( $\log P_{o/w}$ );  $\log P$  väärtused teistes solvendipaarides on käesolevas töös kasutatud valimist saadaval vaid mõne üksiku aine jaoks. Andmete võrdlust kajastab Tabel 2. Tulemuste kokkulangevusele hinnangu andmiseks kasutati saadud väärtuste ruutkeskmist hälvet kirjanduse väärtustest ja vastavate hälvete aritmeetilist keskmist (vabadusastmete arv mõlemal juhul 22). Ruutkeskmiseks hälbeks saadi 0,10 log ühikut. Sellest võib järeldada, et saadud andmed on kirjanduslikele piisavalt lähedased. Hälvete aritmeetiline keskmine on 0,03 log ühikut standardhälbega 0,11 ühikut, millest saab järeldada, et puudub süstemaatiline erinevus eksperimentaalsete ja kirjanduslike andmete vahel. Kuna sama eksperimendimetoodikat kasutati kõikide lahustipaaride korral, siis kokkuvõttes võib lugeda kogutud andmed usaldusväärseteks.

**Tabel 2.** Eksperimentaalsete ja kirjanduslike oktanool/vesi jaotuskoefitsientide võrdlus. Tähistused: Eksp. – eksperimentaalne, Kirj. – kirjanduslik, Erinevus – eksperimendiandmete ja kirjanduslike andmete vahe.

Uuritav aine <sup>a</sup>	$\log P_{o/w}$		
	Eksp.	Kirj. <sup>b</sup>	Erinevus
Furfurool	0,28	0,41	-0,13
Fenool	1,50	1,48 <sup>c</sup>	0,02
2-Isopropüülfenool	2,92	2,88	0,04
3-Metüülfenool	2,00	1,96	0,04
Katehhool	0,92	0,88	0,04
4-Metüülkatehhool	1,36	1,37	-0,01
Resortsinool	0,81	0,80	0,01
1,3-Dimetoksübenseen	2,17	2,21	-0,04
Guajakool	1,30	1,32 <sup>d</sup>	-0,02
Vanilliin	1,20	1,21	-0,01
Isovanilliin	1,04	0,97	0,07
Hüdrokinoon	0,50	0,59	-0,09
4-Aniisaldehyüd	1,57	1,61	-0,04
1,2-Dimetüülbenseen	2,89	3,12	-0,23
Atsetofenoon	1,60	1,58	0,02
2'-Hüdroksüatsetofenoon	1,88	1,92	-0,04
Antrakinoon	3,33	3,39	-0,06
Süringool	1,08	1,15	-0,07
Sirelhape	1,21	1,04	0,17
Gallushape	0,38	0,70	-0,32
Vanilliinhape	1,40	1,43	-0,03
4-Hüdroksübensoehape	1,55	1,58 <sup>c</sup>	-0,03
4-Aniishape	1,95	1,96 <sup>c</sup>	-0,01

<sup>a</sup> Tabelis on toodud ainult soluudid, mille jaoks leidis kirjandusest usaldusväärne  $\log P_{o/w}$  väärtus.

<sup>b</sup> Andmed allikast [33], kui pole teisiti märgitud. <sup>c</sup> Viide [6]; <sup>d</sup> Viide [34].

### 5.3 Hinnang faaside tasakaaluliste koostiste ennustamisele

COSMO-RS meetodiga arvutatud faaside tasakaaluliste koostiste täpsuse hindamiseks võrreldi saadud massiprotsente vastavate eksperimentaalsete andmetega. Tulemusi kajastab Tabel 3.

**Tabel 3.** COSMO-RS meetodiga arvutatud faaside tasakaaluliste koostiste võrdlus eksperimentaalsete andmetega (kirjandusest) vastava orgaanilise solvendi ja vee vastastikuse lahustuvuse kohta (esitatuna g/100g lahuse kohta). Tähistused: T – temperatuur (°C), Eksp. – eksperimentiandmed, Arv. – COSMO-RS meetodiga arvutatud väärtused, Δ% - suhteline ennustusviga (protsentides).

Orgaaniline solvent	Solvendi massiprotsent vees				Vee massiprotsent solvendis			
	T	Eksp. <sup>a</sup>	Arv. <sup>a</sup>	Δ% <sup>b</sup>	T	Eksp. <sup>a</sup>	Arv. <sup>a</sup>	Δ% <sup>b</sup>
<i>n</i> -Heksaan	25	0,00123	0,0034	179%	20	0,0111	0,0032	-71%
Tolueen	25	0,0515	0,050	-3%	25	0,0334	0,021	-37%
Klorobenseen	30	0,0488	0,078	59%	25	0,0327	0,017	-48%
Diklorometaan	25	1,3	3,3	151%	25	0,198	0,078	-61%
MTBE	25	5,166	4,1	-21%	25	2,948	0,32	-89%
Isopropüül-atsetaat	20	2,9	3,4	16%	24	1,90 <sup>c</sup>	4,0	111%
Bensüülalkohol	20	0,08	3,6	4430%	20	8,37	8,7	4%
Tsükloheksanool	24,6	3,75	6,6	75%	20	11.78	20	66%
<i>n</i> -Butanool	25	7,45	0,67	-91%	25	20.50	6,2	-70%
<i>n</i> -Oktanool	25	0,054	0,082	53%	22,52	4,871 <sup>d</sup>	3,2	-35%

<sup>a</sup> Eksperimentaalsed väärtused on allikast [35], kui pole teisiti märgitud. Kõik eksperimentaalsed väärtused on esitatud sama täpsusega, nagu nad on toodud vastavas kirjandusallikas. Arvutuslikud andmed on esitatud kahe tüvenumbriga. <sup>b</sup> Suhteline erinevus arvutusliku ja eksperimentaalse väärtuse vahel. <sup>c</sup> Väärtus on saadud meie rühma varasema uurimistöö käigus. <sup>d</sup> Viide [36].

On oodata, et vead orgaanilise faasi veesisalduse ennustamises mõjutavad lõpptulemust rohkem kui valesti arvutatud veefaasi koostis, kuna vähepolarse keskkonna omadused sõltuvad tugevasti polaarsete lisandite sisaldusest, samas kui vee omadusi mõningane orgaanilise aine sisaldus märkimisväärselt ei mõjuta. Tulemustest järeldub, et COSMO-RS meetod ennustab erinevate omadustega solventide korral veega või küllastamisel tekkivate faaside koostiseid üsnagi sarnaselt. Ennustusvõime kvaliteet on halvem solventide korral, millel on vesiniksideme andmise võime. Ainult vesiniksideme aktseptoorse võimega solventide lahustuvus vees on ennustatud hästi, kuid vee lahustuvus nendes suure veega. Samas *n*-oktanooli, mis võib olla nii vesiniksideme

doonor kui aktseptor, ja vee vastastikusel küllastamisel tekkivate faaside koostiseid ennustab COSMO-RS meetod üsna hästi. Kuna oktanool-vesi jaotumine on populaarne uurimisvaldkond (näiteks toksikoloogias ja ravimite arenduses) ning eksperimentaalseid  $\log P$  väärtusi oktanool-vesi süsteemis on kirjandusest leida märksa rohkem kui teiste solventide puhul, siis on võimalik, et COSMO-RS meetodi parametrisatsioonis kasutatud vastavad parameetrid on täpsemad ja seetõttu ka ennustusvõime võrreldes teiste lahustitega parem.

#### **5.4 Arvutatud $\log D$ väärtuste üldine kokkulangevus eksperimentaalsete andmetega**

Andmete kokkulangevust COSMO-RS arvutuste ja eksperimentaalsete väärtuste vahel hinnati korrelatsioonianalüüsi abil vastavalt võrrandile 6:

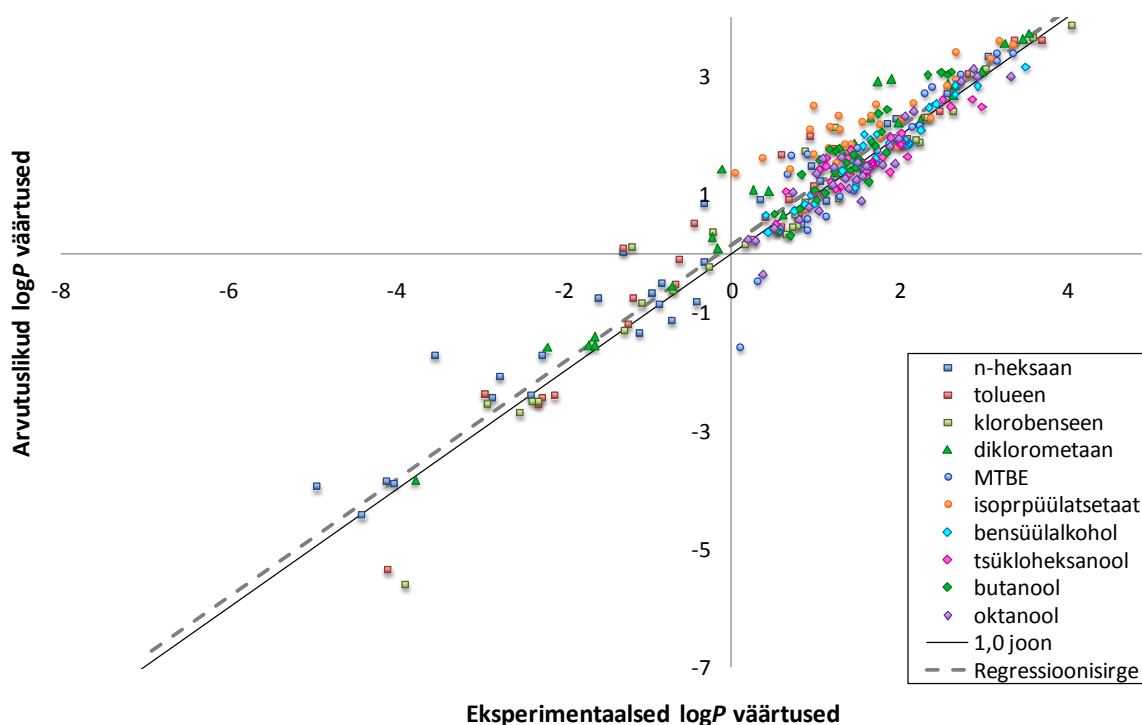
$$\log P(\text{COSMO-RS}) = a \cdot \log P(\text{eksp.}) + b \quad (6)$$

Seejuures kasutati  $\log P$  väärtusi, kuna vastavate eksperimendiandmete kvaliteet oli mõnevõrra kõrgem kui  $\log D$  väärtuste puhul (vähem ekstrapoleeritud väärtusi). Tähtsamad hinnatavad parameetrid on korrelatsiooni standardhälve ( $S$ ) ja korrelatsioonisirge tõus. Lisaks sellele arvutati arvutatud ja eksperimentaalsete ja arvutuslike  $\log P$  väärtuste keskmised erinevused ( $\Delta$ ). Madal regressiooni standardhälve ning ühele lähedane tõus on tähtsad, kui arvutatakse kahe  $\log P$  väärtuse erinevust (nt hindamaks ainete suhtelist lipofiilsust või lahustite suhtelist headust). Olukorras, kus koostatakse empiiriline korrektsioonivõrrand, on oluline saada võimalikult madal standardhälve ning arvutatud väärtuste kasutamisel ilma paranditeta tuleks vaadelda keskmist viga, et hinnata, kas arvutatud väärtused on kalduvad olema süstemaatiliselt üle- või alahinnatud.

Et hinnata, millisel määral ebatäpsused faaside tasakaalulise koostise ennustamises mõjutavad lõpptulemust ( $\log P$  väärtusi), arvutati  $\log P$  väärtused kahel erineval moel: (1) ennustades faaside tasakaalulisi koostisi COSMO-RS meetodiga ja (2) kasutades eksperimendiandmeid faaside tasakaaluliste koostiste kohta. Kuna kirjanduslikke andmeid vastavate lahustite kohta leidis piiratud kogustes, kasutati erinevatel temperatuuridel (20 kuni 30°C, vt Tabel 3) saadud eksperimendiandmeid. Kirjandusest on teada, et temperatuuri efekt lahustuvusele 10°C laiuses temperatuurivahemikus on üsna madal ning sellisest temperatuuride erinevusest tulenev ebatäpsus ei mõjuta oluliselt käesoleva töö tulemusi [36].

Andmete üldist kokkulangevust ennustatud faaside koostiste korral illustreerib Joonis 1. Punktide graafikul vastavad uuritavate ainete  $\log P$  väärtustele vastavas solvendis. Analüüsi

tulemusena saadi korrelatsioonikoefitsiendiks  $R^2 = 0,93$  ning korrelatsiooni standardhälbeks  $S = 0,46$  log ühikut. Tulemuste hindamisel vaadeldi ka tõusu ja vabaliiget, mis olid vastavalt  $0,99(\pm 0,02)$  ja  $0,15(\pm 0,03)$ . Sellest järeldub, et eksperimentaalsete ja ennustatud  $\log P$  väärtuste kokkulangevus on üldjuhul küllalt hea. Eksperimentaalsete faaside koostiste kasutamisel tulemus veidi paraneb ning vastavateks parameetriteks saadi  $R^2 = 0,93$ ,  $S = 0,42$  log ühikut, tõus  $0,96(\pm 0,01)$  ja vabaliige  $0,28(\pm 0,03)$ . Punktide üldine hajuvus jääb praktiliselt samaks.



**Joonis 1. COSMO-RS meetodiga saadud tulemuste ja eksperimentaalsete  $\log P$  väärtuste võrdlus. COSMO-RS arvutustes on kasutatud ennustatud faaside koostiseid.**

Kirjandusest on teada, et COSMO-RS meetodiga saadud tulemused ei lange eksperimendiandmetega täielikult kokku (COSMO-RS väärtused on enamasti kõrgemad), kuid on omavahel korrelatiivses sõltuvuses [23]. Ka käesoleva töö tulemused annavad samasuguse hinnangu.

### **5.5 Arvutuslike ja eksperimentaalsete $\log P$ väärtuste kokkulangevus erinevate orgaaniliste lahustite korral**

Arvutuste kvaliteedi hindamisel lahustite kaupa lähtuti samuti korrelatsioonidest arvutuslike ja eksperimentaalsete  $\log P$  väärtuste vahel ning kasutati samu parameetreid nagu ka üldise kokkulangevuse hindamise korral ( $R^2$  ja  $S$ ).

Tulemused on kajastatud Tabelis 4. Arvutatud faaside koostiste kasutamisel esines kõige parem COSMO-RS arvutuste ja eksperimendiandmete kokkulangevus oktanooli ja klorobenseeni korral (vabaliige nullilähedane, tõus läheneb ühele). Punktide hajuvus (regressiooni standardhälve  $S$ ) on madalaim butanooli, bensüülalkoholi ja tsükloheksanooli korral ning on kõrgem aprotoonsete lahustite puhul. Osaliselt võib see olla põhjustatud alkoholidele vastavate eksperimendiandmete mõnevõrra paremast kvaliteedist võrreldes aprotoonsete lahustitega. Samas on Tabelist 4 näha, et erinevate solventide  $S$  väärtuste varieeruvust ei saa seletada üksnes eksperimentide korratavuse varieeruvusega.

Enamikel juhtudel eksperimentaalsete faaside koostiste kasutamine parandab arvutusmeetodi ennustusvõimet – tõus on lähedasem ühele. Punktide üldine hajuvus ( $S$ ) jääb praktiliselt samaks nii eksperimentaalsete kui ka ennustatud faaside koostiste kasutamisel. Samas eksperimendiandmeid kasutades erineb vabaliige enamasti veidi enam nullist kui ennustatud koostiste korral.

Ilma vesiniksideme donoorse ega aktseptoorse võimeta orgaaniliste lahustite puhul on regressioonide vabaliikmed positiivsed ja tõusud lähedased ühele (0,92-1,03). Protoonsete lahustite puhul on vabaliikmed pigem nullilähedased või positiivsed ja sirge tõusud varieeruvad 0,8-1,2 piires. Ainult vesiniksideme aktseptoorse võimega lahustite, eriti isopropüülatsetaadi, korrelatsiooniparameetrid erinevad üsna tugevasti teoreetilistest väärtustest. Keskmine erinevus ennustatud ja eksperimentaalsete väärtuste vahel on enamasti suurem juhul, kui faaside koostisena on kasutatud eksperimendiandmeid. Samas, mõningate lahustite korral (oktanool, butanool, isopropüülatsetaat) on tulemus vastupidine ehk keskmine viga on eksperimentaalsete faaside koostiste kasutamisel madalam.

**Tabel 4. COSMO-RS tulemuste kokkulangevus eksperimentaalsete andmetega erinevate orgaaniliste lahustite korral. Hinnangu andmiseks on kasutatud lineaarset regressiooni kujul  $\log P(\text{COSMO-RS}) = a \cdot \log P(\text{eksp.}) + b$ . Tähistused:  $S_{\text{eksp}}$  – eksperimendiandmete päevadevaheline korratavus (kogutud standardhälve),  $a$  – tõus,  $b$  – vabaliige,  $s(a)$  – tõusu standardviga,  $s(b)$  – vabaliikme standardviga,  $R^2$  – korrelatsioonikoefitsient,  $S$  – regressiooni standardhälve,  $\Delta$  – keskmine viga (ennustatud ja eksp. väärtuste vahe).**

Orgaaniline solvent	$S_{\text{eksp}}$	Faaside koostis <sup>a</sup>	$a$	$s(a)$	$b$	$s(b)$	$R^2$	$S$	$\Delta$
<i>n</i> -Heksaan	0,26	Ennust.	1,00	0,05	0,27	0,13	0,936	0,63	0,27
		Eksp.	0,98	0,05	0,32	0,12	0,938	0,61	0,34
Tolueen	0,08	Ennust.	1,03	0,05	0,16	0,10	0,940	0,52	0,17
		Eksp.	1,00	0,05	0,23	0,09	0,942	0,49	0,23
Klorobenseen	0,09	Ennust.	1,03	0,05	0,00	0,10	0,945	0,52	0,01
		Eksp.	0,99	0,04	0,11	0,09	0,947	0,49	0,10
Diklorometaan	0,14	Ennust.	1,01	0,04	0,35	0,08	0,960	0,40	0,36
		Eksp.	0,92	0,04	0,69	0,08	0,957	0,38	0,61
Metüül- <i>tert</i> -butüüleeter	0,12	Ennust.	1,20	0,10	-0,34	0,18	0,834	0,47	-0,04
		Eksp.	0,90	0,07	0,50	0,13	0,847	0,34	0,35
Isopropüül-atsetaat	0,04	Ennust.	0,69	0,08	1,08	0,15	0,716	0,36	0,56
		Eksp.	0,77	0,08	0,88	0,15	0,767	0,36	0,50
Bensüül-alkohol	0,05	Ennust.	0,92	0,05	0,14	0,10	0,912	0,21	0,00
		Eksp.	0,95	0,06	0,19	0,11	0,913	0,22	0,10
Tsükloheksaanool	0,05	Ennust.	0,76	0,07	0,35	0,12	0,817	0,22	-0,05
		Eksp.	0,90	0,08	0,23	0,14	0,827	0,25	0,07
<i>n</i> -Butanool	0,06	Ennust.	1,23	0,09	-0,12	0,15	0,860	0,28	0,22
		Eksp.	0,88	0,07	0,23	0,11	0,855	0,21	0,06
<i>n</i> -Oktanool	0,02	Ennust.	1,03	0,07	-0,09	0,12	0,882	0,30	-0,05
		Eksp.	0,98	0,07	0,05	0,11	0,887	0,28	0,02

<sup>a</sup> Ennust. – nii  $\log P$  kui ka faaside tasakaaluline koostis on ennustatud COSMO-RS meetodiga; Eksp. –  $\log P$  on leitud COSMO-RS meetodil, aga faaside tasakaaluline koostis on leitud eksperimendist.

## 5.6 Arvutuslike ja eksperimentaalsete $\log P$ väärtuste kokkulangevuse hinnang erinevate soluutide korral

Võrreldes COSMO-RS arvutuste ja eksperimendiandmete kokkulangevust uuritud ainete kaupa jäid  $R^2$  väärtused vahemikku 0,59-0,98 ning regressiooni standardhälbed vahemikku 0,11-0,89 log ühikut. Kuigi enamiku uuritavate ainete korral langesid arvutustulemused eksperimentaalsetega hästi kokku ning andsid kõrgeid korrelatsioonikoefitsiendi väärtusi, esines ka ühendeid, mille ennustamine ei andnud nii häid tulemusi. Sellisteks ühenditeks on süringool, sirelaldehyüd, sirelhape ja 3-metoksükatehool. Kui võrrelda nende  $\log P(\text{COSMO-RS}) = a \cdot \log P(\text{eksp.}) + b$  korrelatsioone teiste ainete omadega, ilmneb, et sirgete vabaliikmed on ebatavaliselt suured. See tähendab, et vead on süstemaatiliselt ühes suunas ehk kõik ennustatud  $\log P$  väärtused on reaalistest väärtustest tunduvalt kõrgemad. Vaadeldes ühendite struktuure võib oletada, et COSMO-RS meetod alahindab hüdrofiilse(te) rühma(de) vastasmõjude tugevust keskkonnaga, kui nimetatud rühmad saavad moodustada mitut sisemolekulaarset vesiniksidet. Põhjus võib peituda mitte statistilise termodünaamika vaid geomeetria optimeerimise (DFT) protseduuris: saadud konformeeride hulgas pole ühtegi ilma sisemolekulaarse vesiniksidemeta, kuid pole välistatud, et selline (teistest hüdrofiilsem) konformeer võib esineda reaalses lahuses. Tähelepanelik tuleb olla ka ka 1,3-dimetoksübenseeni ja 1,4-dimetoksübenseeni korral, mille ennustatud  $\log P$  väärtused hõlbivad samuti enamasti ühes suunas (on ülehinnatud), kuid tõenäolist põhjust sellele ei tuvastatud.

Mõningates solventides (*n*-heksaan, toluen, klorobenseen, metüül-*tert*-butüüleeter) hõlbis gallushape võrreldes teiste ainetega olulisel määral. See võib olla põhjustatud sellest, et selle aine jaotuskoefitsient on sedavõrd madal, et on kasutatud metoodika jaoks piiri peal ja on määratud keskmisest suurema veaga.



## 6 Kokkuvõte

Käesoleva töö käigus uuriti eksperimentaalselt ligniini derivaatide ning lähedaste ühendite jaotuskoefitsiente 20 erinevas solvendipaaris ning hinnati saadud andmete abil COSMO-RS meetodi ennustuvõimet.

Koostatud andmestiku täielikkus (andmed on olemas kõigi uuritud ühendite jaoks kõigis solvendipaarides) on väga väärtuslik seetõttu, et võimaldab andmematriksit edaspidi mitmekülgselt kasutada. Näiteks saab andmematriksile rakendada vabalt valitavaid statistilisi meetodeid (nt PCA), kasutada jaotuskoefitsientide ennustamise mudelite arendamiseks või valideerimiseks, samuti jaotuskoefitsientide mõõtemetoodikate arendamiseks.

Töö käigus leiti, et COSMO-RS meetodi võime ennustada jaotuskoefitsiente sõltub uuritavast ainest ning kasutatavast lahustist. Ennustused on üldiselt täpsemad lahustite korral, mis on nii vesiniksideme doonorid kui aktseptorid (bensüülalkohol, butanool, oktanool, tsükloheksanool) ja vähem täpsed aprotoonsete lahustite puhul. Teisalt tuleb COSMO-RS meetodiga töötades olla ettevaatlik kasutades aineid, kus polaarsel asendusrühmal on mitu võimalust moodustada sisemolekulaarset vesiniksidet (näited käesolevast tööst – süringool, sirelaldehyüd, sirelhape ja 3-metoksükatehhool). Nimetatud ainete ennustatud  $\log P$  väärtused on eksperimentaalsetest tugevalt ja süstemaatiliselt kõrgemad. Üldjuhul on eksperimentaalsete ja arvutuslike andmete kokkulangevus hea ja väärtused on omavahel tugevas korrelatiivses sõltuvuses. See kinnitab kaudselt ka loodud eksperimentaalse andmestiku usaldusväärsust.

## 7 Summary

### **Experimental study and COSMO-RS modelling of partition coefficients of lignin decomposition products**

In this thesis partition/distribution coefficients of lignin derivatives and other similar compounds in 20 different solvent pairs were investigated experimentally. The obtained data were used to evaluate prediction capability of the COSMO-RS method.

The obtained data matrix is complete, i.e. without gaps, which is very valuable by allowing the data to be used in many ways. Variety of statistical methods can be applied to data matrix, for example PCA. Data matrix can also be used for developing or validating partition/distribution coefficient prediction models. Besides, the data can be used to develop partition/distribution coefficient measurement methods.

This thesis shows that the capability of COSMO-RS method to predict partition/distribution coefficients depends on the substance of interest and used solvents. Predictions are generally better when using solvents that are simultaneously donors and acceptors of hydrogen bond (benzyl alcohol, butanol, octanol, cyclohexanol). Predictions are less accurate with aprotic solvents. On the other hand, when working with the COSMO-RS method one has to be careful if using substances where polar substituents have several possibilities of forming intermolecular hydrogen bonds. Examples from this work are syringol, syringaldehyde, syringic acid and 3-methoxycatechol. The predicted  $\log P$  values of these substance are systematically higher than experimental values. In general the agreement between experimental and computational data is good and values are strongly correlated. This also indirectly confirms the reliability of the obtained experimental data.

## Kasutatud kirjandus

1. Mu, W.; Ben, H.; Ragauskas, A.; Deng, Y. Lignin Pyrolysis Components and Upgrading—Technology Review. *Bioenerg. Res.* **2013**, *6* (4), 1183–1204.
2. Long, J.; Xu, Y.; Wang, T.; Shu, R.; Zhang, Q.; Zhang, X.; Fu, J.; Ma, L. Hydrothermal Depolymerization of Lignin: Understanding the Structural Evolution. *BioResources* **2014**, *9* (4), 7162–7175.
3. Putro, J. N.; Soetaredjo, F. E.; Lin, S.-Y.; Ju, Y.-H.; Ismadji, S. Pretreatment and Conversion of Lignocellulose Biomass into Valuable Chemicals. *RSC Adv.* **2016**, *6* (52), 46834–46852.
4. Kleinert, M.; Barth, T. Phenols from Lignin. *Chem. Eng. Technol.* **2008**, *31* (5), 736–745.
5. Vigneault, A.; Johnson, D. K.; Chornet, E. Base-Catalyzed Depolymerization of Lignin: Separation of Monomers. *Can. J. Chem. Eng.* **2007**, *85* (6), 906–916.
6. Leo, A.; Hansch, C.; Elkins, D. Partition Coefficients and Their Uses. *Chem. Rev.* **1971**, *71* (6), 525–616.
7. Berthod, A.; Carda-Broch, S. Determination of Liquid–liquid Partition Coefficients by Separation Methods. *J. Chromatogr. A* **2004**, *1037* (1–2), 3–14.
8. Skoog, D. A.; West, D. M.; Holler, J. F.; Crouch, S. R. *Skoog and West's Fundamentals of Analytical Chemistry*, 9th edition, Cengage technology edition.; Cengage Learning: Hampshire, 2014.
9. Cherkasov, A.; Muratov, E. N.; Fourches, D.; Varnek, A.; Baskin, I. I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y. C.; Todeschini, R.; et al. QSAR Modeling: Where Have You Been? Where Are You Going To? *J. Med. Chem.* **2014**, *57* (12), 1–72.
10. Rinaldi, R.; Jastrzebski, R.; Clough, M. T.; Ralph, J.; Kennema, M.; Bruijninx, P. C. A. B.; Weckhuysen, B. M. Paving the Way for Lignin Valorisation: Recent Advances in Bioengineering, Biorefining and Catalysis. *Angew. Chem., Int. Ed* **2016**, *55* (29), 8164–8215.
11. Hopkins, A. L.; Kserü, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H. The Role of Ligand Efficiency Measures in Drug Discovery. *Nat. Rev. Drug Discovery* **2014**, *13* (2), 105–121.
12. Leeson, P. D.; Springthorpe, B. The Influence of Drug-like Concepts on Decision-Making in Medicinal Chemistry. *Nat. Rev. Drug Discovery* **2007**, *6* (11), 881–890.

13. Rumynin, V. G.; Nikulenkov, A. M. Geological and Physicochemical Controls of the Spatial Distribution of Partition Coefficients for Radionuclides (Sr-90, Cs-137, Co-60, Pu-239,240 and Am-241) at a Site of Nuclear Reactors and Radioactive Waste Disposal (St. Petersburg Region, Russian Federation). *J. Environ. Radioact.* **2016**, *162–163*, 205–218.
14. Pan, J.; Fu, J.; Deng, S.; Lu, X. Distribution Coefficient of Products from Lignin Oxidative Degradation in Organic-Water Systems. *Fuel Process. Technol.* **2015**, *140*, 262–266.
15. Bao, J. J.; Liu, X.; Zhang, Y.; Li, Y. The Development of a High-Throughput Measurement Method of Octanol/Water Distribution Coefficient Based on Hollow Fiber Membrane Solvent Microextraction Technique. *J. Chromatogr. B* **2014**, *967*, 183–189.
16. Tropsha, A. Best Practices for QSAR Model Development, Validation, and Exploitation. *Mol. Inf.* **2010**, *29* (6–7), 476–488.
17. Gramatica, P. Principles of QSAR Models Validation: Internal and External. *QSAR Comb. Sci.* **2007**, *26* (5), 694–701.
18. Klamt, A. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* **1995**, *99* (7), 2224–2235.
19. Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J. C. W. Refinement and Parametrization of COSMO-RS. *J. Phys. Chem. A* **1998**, *102* (26), 5074–5085.
20. Klamt, A.; Eckert, F.; Arlt, W. COSMO-RS: An Alternative to Simulation for Calculating Thermodynamic Properties of Liquid Mixtures. *Annu. Rev. Chem. Biomol. Eng.* **2010**, *1*, 101–122.
21. Eckert, F.; Klamt, A. Validation of the COSMO-RS Method: Six Binary Systems. *Ind. Eng. Chem. Res.* **2001**, *40* (10), 2371–2378.
22. Buggert, M.; Cadena, C.; Mokrushina, L.; Smirnova, I.; Maginn, E. J.; Arlt, W. COSMO-RS Calculations of Partition Coefficients: Different Tools for Conformation Search. *Chem. Eng. Technol.* **2009**, *32* (6), 977–986.
23. Klamt, A.; Eckert, F.; Reinisch, J.; Wichmann, K. Prediction of Cyclohexane-Water Distribution Coefficients with COSMO-RS on the SAMPL5 Data Set. *J. Comput.-Aided Mol. Des.* **2016**, *30* (11), 959–967.
24. Tšepelevitš, S. Arvutuslik Mudel Vedelik-Vedelik Ekstraksiooni Tulemuste Ennustamiseks. Magistritöö, Tartu, 2012.

25. Noubigh, A.; Mgaidi, A.; Abderrabba, M. Temperature Effect on the Distribution of Some Phenolic Compounds: An Experimental Measurement of 1-Octanol/Water Partition Coefficients. *J. Chem. Eng. Data* **2010**, *55* (1), 488–491.
26. Tshepelevitsh, S.; Trummal, A.; Haav, K.; Martin, K.; Leito, I. Hydrogen-Bond Donicity in DMSO and Gas Phase and Its Dependence on Brønsted Acidity. *J. Phys. Chem. A* **2017**, *121* (1), 357–369.
27. Eckert, F.; Klamt, A. Fast Solvent Screening via Quantum Chemistry: COSMO-RS Approach. *AIChE J.* **2002**, *48* (2), 369–385.
28. F. Eckert and A. Klamt, COSMOtherm, Version C3.0, Release 17.01; COSMOlogic GmbH & Co. KG. <http://www.cosmologic.de>.
29. Abboud, J.-L. M.; Notari, R. Critical Compilation of Scales of Solvent Parameters. Part I. Pure, Non-Hydrogen Bond Donor Solvents. *Pure Appl. Chem., PAC* **1999**, *71* (4), 645–718.
30. Reichardt, C. Empirical Parameters of Solvent Polarity as Linear Free-Energy Relationships. *Angew. Chem., Int. Ed. Engl.* **1979**, *18* (2), 98–110.
31. Shakerizadeh-Shirazi, F.; Hemmateenejad, B.; Mehranpour, A. M. Determination of the Empirical Solvent Polarity Parameter ET(30) by Multivariate Image Analysis. *Anal. Methods* **2013**, *5* (4), 891–896.
32. Reichardt, C. Solvatochromic Dyes as Solvent Polarity Indicators. *Chem. Rev.* **1994**, *94* (8), 2319–2358.
33. Hansch, C.; Leo, A.; Hoekman, D. *Exploring QSAR: Hydrophobic, Electronic, and Steric Constants*; American Chemical Society: Washington DC, 1995.
34. Zissimos, A. M.; Abraham, M. H.; Barker, M. C.; Box, K. J.; Tam, K. Y. Calculation of Abraham Descriptors from Solvent–Water Partition Coefficients in Four Different Systems; Evaluation of Different Methods of Calculation. *J. Chem. Soc., Perkin Trans. 2* **2002**, No. 3, 470–477.
35. Riddick, J. A.; Bunger, W. B.; Sakano, T. K. *Organic Solvents: Physical Properties and Methods of Purification, 4th Edition*; John Wiley & Sons, Inc., 1986; Vol. 2.
36. Lang, B. E. Solubility of Water in Octan-1-ol from (275 to 369) K. *J. Chem. Eng. Data* **2012**, *57* (8), 2221–2226.

## Lisad

### Lisa 1. Eksperimentaalselt leitud log*D* väärtused

**Tähistused:** nHex – *n*-heksaan; Tol – toluen; PhCl – klorobenseen; DCM – diklorometaan; MTBE – metüül-*tert*-butüüleeter; iPrAc – isopropüülatsetaat; Ben.alc. – bensüülalkohol; cHexOH – tsükloheksanool; BuOH – butanool; Oct – oktanool.

Soluut	veefaas - vesi/fosfaatpuhver pH≈7									
	nHex	Tol	PhCl	DCM	MTBE	iPrAc	Ben.alc.	cHexOH	BuOH	Oct
Furfurool	-0,42	0,59	0,78	1,37	0,56	0,88	0,99	0,51	0,53	0,27
Fenool	-0,96	0,25	0,17	0,62	1,74	1,77	1,46	1,68	1,44	1,50
2-Isopropüülfenool	1,06	2,08	2,02	2,26	3,22	3,21	2,64	2,95	2,62	2,89
3-Metüülfenool	-0,34	0,80	0,74	1,21	2,13	2,18	1,85	2,02	1,79	1,99
Katehool	-2,85	-1,24	-1,34	-0,72	1,11	1,10	0,89	1,16	0,94	0,86
4-Metüülkatehool	-2,23	-0,68	-0,74	-0,14	1,49	1,52	1,30	1,49	1,34	1,35
Resortsinool	-4,09	-2,18	-2,36	-1,63	0,84	0,95	0,89	1,28	1,00	0,84
1,3-Dimetoksübenseen	2,10	2,84	2,96	3,36	2,48	2,72	2,39	2,04	1,88	2,16
Guajakool	0,33	1,17	1,33	1,69	1,47	1,65	1,51	1,38	1,25	1,30
Vanilliin	-1,01	0,53	0,73	1,30	0,97	1,29	1,60	1,28	1,14	1,07
Isovanilliin	-1,11	0,40	0,70	1,33	0,89	1,24	1,59	1,24	1,07	1,03
Koniferüülaldehüüd	-0,93	0,93	1,24	1,93	1,27	1,78	2,09	1,65	1,50	1,47
Hüdrokinoon	-4,18	-2,27	-2,49	-1,66	0,55	0,69	0,78	1,08	0,82	0,53
1,4-dimetoksübenseen	1,90	2,67	2,88	3,27	2,32	2,56	2,35	1,92	1,72	1,98
4-Aniisalaldehüüd	0,84	2,04	2,22	2,63	1,65	1,86	2,13	1,59	1,51	1,55
1,2-Dimetüülbenseen	3,15	3,38	3,53	3,56	3,10	3,07	2,80	2,67	2,47	2,89
Atsetofenoon	1,12	2,03	2,20	2,68	1,75	1,98	2,03	1,61	1,52	1,60
2'-Hüdroksüatsetofenoon	1,56	2,43	2,62	2,94	2,22	2,39	2,23	1,78	1,69	1,87
Antrakinoon	2,63	3,65	4,08	3,83	3,37	3,48	3,55	2,94	2,61	3,29
9-Hüdroksüfluoreen	0,95	2,30	2,33	2,60	2,64	2,70	2,86	2,56	2,33	2,76
Pürogallool	-5,00	-2,91	-3,04	-2,12	0,33	0,41	0,42	0,69	0,54	0,21
3-Metoksükatehool	-1,63	-0,50	-0,31	0,23	0,70	0,92	0,98	0,97	0,84	0,66
Süringool	-0,32	0,93	1,24	1,88	0,91	1,28	1,61	1,12	1,16	1,07
Sirelalaldehüüd	-1,46	0,38	0,71	1,56	0,50	1,01	1,63	0,98	0,96	0,86
Sirelhape	-6,19	-3,95	-3,84	-2,77	-1,95	-1,68	-1,01	-1,23	-0,81	-1,45
Gallushape	-9,07	-6,76	-6,57	-6,44	-2,57	-2,62	-2,24	-1,67	-1,80	-2,30
Vanilliinhape	-6,03	-4,45	-4,32	-3,07	-1,45	-1,26	-0,80	-0,78	-0,57	-1,14
4-Hüdroksübensoehape	-6,83	-4,73	-4,95	-4,05	-0,97	-1,07	-0,95	-0,53	-0,50	-0,87
Trans-feruulhape	-5,21	-3,06	-2,71	-1,99	-0,92	-0,72	-0,37	-0,42	-0,09	-0,65
4-Aniishape	-4,04	-2,44	-2,32	-1,65	-0,70	-0,59	-0,50	-0,49	-0,50	-0,61

Soluut	veefaas – 0,1M HCl									
	nHex	Tol	PhCl	DCM	MTBE	iPrAc	Ben.alc.	cHexOH	BuOH	Oct
Furfurool	-0,42	0,60	0,79	1,37	0,58	0,88	0,99	0,53	0,54	0,28
Fenool	-0,95	0,26	0,17	0,62	1,73	1,77	1,47	1,69	1,45	1,50
2-Isopropüülfenool	1,05	2,10	2,00	2,26	3,17	3,20	2,67	2,87	2,61	2,92
3-Metüülfenool	-0,33	0,81	0,75	1,21	2,16	2,17	1,88	2,01	1,78	2,00
Katehhool	-2,85	-1,23	-1,28	-0,70	1,13	1,16	0,95	1,19	0,98	0,92
4-Metüülkatehhool	-2,25	-0,67	-0,69	-0,16	1,49	1,56	1,32	1,55	1,37	1,36
Resortsinool	-4,03	-2,12	-2,30	-1,63	0,90	0,99	0,87	1,31	1,01	0,81
1,3-Dimetoksübenseen	1,96	2,82	3,03	3,48	2,39	2,68	2,44	2,03	1,85	2,17
Guajakool	0,34	1,21	1,37	1,67	1,50	1,67	1,56	1,39	1,27	1,30
Vanilliin	-0,82	0,69	0,87	1,47	1,12	1,43	1,72	1,41	1,23	1,20
Isovanilliin	-1,10	0,40	0,73	1,35	0,90	1,24	1,60	1,25	1,10	1,04
Koniferüülaldehüüd	-0,87	0,98	1,27	1,99	1,30	1,77	2,11	1,69	1,53	1,48
Hüdrokinoon	-4,11	-2,26	-2,38	-1,70	0,56	0,69	0,74	1,10	0,81	0,50
1,4-dimetoksübenseen	1,86	2,72	2,88	3,26	2,30	2,57	2,35	1,90	1,74	2,06
4-Aniisalaldehüüd	0,83	2,06	2,24	2,67	1,65	1,90	2,13	1,61	1,47	1,57
1,2-Dimetüülbenseen	3,07	3,38	3,59	3,55	3,16	3,10	2,66	2,61	2,49	2,89
Atsetofenoon	1,13	2,05	2,20	2,65	1,73	1,97	2,06	1,61	1,51	1,60
2'-Hüdroksüatsetofenoon	1,58	2,48	2,65	2,99	2,24	2,38	2,25	1,80	1,67	1,88
Antrakinoon	2,58	3,71	4,05	3,82	3,36	3,36	3,50	2,99	2,57	3,33
9-Hüdroksüfluoreen	0,95	2,31	2,34	2,59	2,73	2,68	2,93	2,52	2,34	2,74
Pürogallool	-4,95	-2,94	-2,91	-2,19	0,31	0,37	0,40	0,65	0,50	0,19
3-Metoksükatehhool	-1,59	-0,45	-0,23	0,27	0,76	0,94	1,05	1,05	0,83	0,72
Süringool	-0,32	0,94	1,24	1,91	0,91	1,28	1,58	1,13	1,16	1,08
Sirelalaldehüüd	-1,30	0,60	0,88	1,75	0,67	1,18	1,73	1,13	1,04	1,01
Sirelhape	-3,53	-1,29	-1,18	-0,11	0,71	0,98	1,65	1,43	1,24	1,21
Gallushape	-6,40	-4,09	-3,90	-3,76	0,10	0,05	0,43	1,00	0,69	0,38
Vanilliinhape	-2,39	-1,17	-1,07	-0,23	1,19	1,29	1,63	1,68	1,41	1,40
4-Hüdroksübensoehape	-4,41	-2,31	-2,53	-1,63	1,45	1,36	1,47	1,89	1,63	1,55
Trans-feruulhape	-2,77	-0,62	-0,27	0,45	1,52	1,73	2,07	2,02	1,71	1,79
4-Aniishape	-0,71	0,54	0,65	1,24	1,94	2,04	2,10	2,10	1,84	1,95

## Lisa 2. COSMO-RS meetodiga arvutatud log*P* väärtused

**Tähistused:** nHex – n-heksaan; Tol – toluen; PhCl – klorobenseen; DCM – diklorometaan; MTBE – metüül-*tert*-butüüleeter; iPrAc – isopropüülatsetaat; Ben.alc. – bensüülalkohol; cHexOH – tsükloheksanool; BuOH – butanool; Oct – oktanool.

Soluut	Faaside koostised ennustatud COSMO-RS meetodiga									
	nHex	Tol	PhCl	DCM	MTBE	iPrAc	Ben.alc.	cHexOH	BuOH	Oct
Furfurool	-0,82	0,46	0,47	1,41	0,44	0,77	0,85	0,50	0,40	0,21
Fenool	-0,66	0,23	0,16	0,65	1,77	2,20	1,55	1,55	1,69	1,55
2-Isopropüülfenool	1,23	1,95	1,90	2,25	3,28	3,61	2,71	2,61	3,09	3,01
3-Metüülfenool	-0,14	0,69	0,66	1,15	2,15	2,57	1,89	1,84	2,07	1,94
Katehool	-2,44	-1,19	-1,30	-0,56	0,65	1,80	1,00	1,21	1,08	0,79
4-Metüülkatehool	-1,71	-0,51	-0,62	0,09	1,19	2,25	1,41	1,56	1,54	1,28
Resortsinool	-3,89	-2,41	-2,49	-1,55	0,41	1,69	0,84	1,13	0,92	0,59
1,3-Dimetoksübenseen	2,29	3,04	3,14	3,64	2,83	2,96	2,54	2,03	2,45	2,42
Guajakool	0,92	1,78	1,79	2,30	1,80	2,35	1,83	1,61	1,78	1,64
Vanilliin	-0,49	0,92	0,88	1,85	1,04	1,87	1,62	1,35	1,40	1,17
Isovanilliin	-1,33	0,61	0,45	1,66	0,60	1,56	1,39	1,10	1,03	0,74
Koniferüülaldehyd	-0,85	1,15	0,95	2,22	0,97	2,00	1,88	1,49	1,54	1,25
Hüdrokinoon	-3,86	-2,44	-2,50	-1,57	0,39	1,44	0,73	0,99	0,76	0,43
1,4-dimetoksübenseen	2,20	2,93	3,04	3,56	2,73	2,86	2,48	1,97	2,38	2,34
4-Aniisaldehyd	0,50	1,92	1,89	2,86	1,52	1,84	1,86	1,30	1,46	1,31
1,2-Dimetüülbenseen	3,34	3,61	3,65	3,73	3,40	3,32	2,85	2,49	3,08	3,12
Atsetofenoon	0,90	1,84	1,93	2,67	1,55	1,82	1,87	1,39	1,60	1,49
2'-Hüdroksiätssetofenoon	1,42	2,41	2,42	3,08	2,18	2,32	2,10	1,63	1,89	1,81
Antrakinoon	2,70	3,62	3,87	4,64	3,40	3,55	3,17	2,48	3,04	3,01
9-Hüdroksüfluoreen	1,49	2,32	2,32	2,87	3,05	3,44	2,85	2,60	3,03	2,93
Pürogallool	-3,94	-2,38	-2,54	-1,59	-0,45	1,63	0,66	1,05	0,68	0,24
3-Metoksükatehool	-0,74	0,51	0,38	1,08	0,74	2,11	1,43	1,42	1,35	1,05
Süringool	0,85	2,00	2,14	2,96	1,69	2,36	2,03	1,61	1,77	1,61
Sirelaldehyd	0,04	1,68	1,74	2,92	1,35	2,16	2,02	1,48	1,59	1,37
Sirelhape	-1,72	0,10	0,12	1,43	1,68	2,52	1,94	1,75	1,74	1,47
Gallushape	-7,97	-5,35	-5,59	-3,82	-1,58	1,38	0,37	1,04	0,31	-0,35
Vanilliinhape	-2,39	-0,75	-0,84	0,28	1,23	2,13	1,46	1,53	1,42	1,11
4-Hüdroksübensoehape	-4,43	-2,53	-2,70	-1,42	1,09	1,86	1,12	1,38	1,22	0,90
Trans-feruulhape	-2,08	-0,09	-0,23	1,06	1,48	2,55	1,93	1,87	1,83	1,51
4-Aniishape	-1,13	0,38	0,33	1,37	1,97	2,30	1,84	1,64	1,72	1,53



Soluut	Eksperimendiandmed faaside koostise kohta									
	nHex	Tol	PhCl	DCM	MTBE	iPrAc	Ben.alc.	cHexOH	BuOH	Oct
Furfurool	-0,77	0,48	0,50	1,49	0,60	0,74	0,91	0,48	0,49	0,28
Fenool	-0,59	0,30	0,26	0,92	1,98	2,18	1,64	1,72	1,43	1,57
2-Isopropüülfenool	1,28	2,00	1,98	2,51	3,45	3,62	2,85	2,94	2,51	3,01
3-Metüülfenool	-0,09	0,76	0,76	1,41	2,35	2,56	1,99	2,05	1,73	1,97
Katehhool	-2,32	-1,07	-1,13	-0,19	1,22	1,70	1,07	1,28	1,04	0,87
4-Metüülkatehhool	-1,61	-0,41	-0,47	0,44	1,72	2,17	1,50	1,68	1,40	1,35
Resortsinool	-3,77	-2,27	-2,29	-1,05	1,00	1,59	0,91	1,19	0,92	0,67
1,3-Dimetoksübenseen	2,29	3,04	3,14	3,71	2,81	3,02	2,68	2,27	2,02	2,42
Guajakool	0,91	1,78	1,80	2,38	2,02	2,34	1,93	1,76	1,54	1,68
Vanilliin	-0,43	0,95	0,92	2,01	1,49	1,77	1,72	1,42	1,32	1,25
Isovanilliin	-1,25	0,65	0,51	1,81	1,08	1,45	1,48	1,12	1,06	0,84
Koniferüülaldehüüd	-0,75	1,19	1,01	2,42	1,61	1,85	2,00	1,55	1,48	1,36
Hüdrokinoon	-3,77	-2,33	-2,34	-1,12	0,91	1,34	0,79	1,03	0,81	0,52
1,4-dimetoksübenseen	2,20	2,93	3,05	3,64	2,72	2,91	2,61	2,19	1,97	2,35
4-Aniisalaldehüüd	0,56	1,94	1,92	2,97	1,65	1,82	1,97	1,39	1,32	1,37
1,2-Dimetüülbenseen	3,34	3,61	3,66	3,79	3,34	3,39	2,99	2,81	2,47	3,10
Atsetofenoon	0,94	1,86	1,95	2,78	1,68	1,79	1,98	1,48	1,44	1,54
2'-Hüdroksüatsetofenoon	1,43	2,42	2,43	3,17	2,21	2,35	2,22	1,79	1,63	1,84
Antrakinoon	2,72	3,63	3,88	4,75	3,41	3,62	3,34	2,77	2,51	3,02
9-Hüdroksüfluoreen	1,51	2,34	2,35	3,04	3,28	3,43	3,01	2,87	2,55	2,96
Pürogallool	-3,83	-2,26	-2,36	-1,14	0,65	1,43	0,73	1,04	0,82	0,37
3-Metoksükatehhool	-0,75	0,53	0,40	1,25	1,42	2,01	1,53	1,49	1,30	1,15
Süringool	0,88	2,02	2,15	3,05	1,93	2,34	2,15	1,72	1,60	1,68
Sirelaldehüüd	0,10	1,71	1,78	3,06	1,72	2,10	2,15	1,56	1,48	1,45
Sirelhape	-1,63	0,18	0,23	1,75	2,13	2,45	2,06	1,87	1,65	1,57
Gallushape	-7,66	-5,07	-5,19	-2,98	0,11	1,04	0,43	0,93	0,72	-0,11
Vanilliinhape	-2,30	-0,66	-0,72	0,62	1,77	2,02	1,57	1,62	1,37	1,22
4-Hüdroksübensoehape	-4,16	-2,32	-2,40	-0,81	1,59	1,75	1,21	1,47	1,19	1,00
Trans-feruulhape	-1,98	-0,01	-0,12	1,38	2,11	2,42	2,06	1,99	1,73	1,62
4-Aniishape	-1,04	0,46	0,44	1,66	2,19	2,28	1,95	1,78	1,55	1,59

### Lisa 3. Kasutatud kemikaalide päritolu ja puhtus

Nimetus	Tootja/Päritolu	Puhtus
Furfurool	Aldrich	99%
Fenool	Riedel-de-Haën	99.5%
2-Isopropüülfenool	ABCR	98%
3-Metüülfenool	Reakhim	„puhas“
Katehhool	Reakhim	„puhas“
4-Metüülkatehhool	Aldrich	≥95%
Resortsinool	Schering-Kahlbaum AG	p. A.
1,3-Dimetoksübenseen	Aldrich	≥98%
Guajakool	Reakhim	„puhas“
Vanilliin	Alfa Aesar	99%
Isovanilliin	Fluka	≥99.5%
Koniferüülaldehüüd	Sigma-Aldrich	98%
Hüdrokinoon	Reakhim	„kõrgem sort“
1,4-dimetoksübenseen	Aldrich	99%
4-Aniisalaldehüüd	Sigma	pole määratletud
1,2-Dimetüülbenseen	Reakhim	„puhas“
Atsetofenoon	Reakhim	„puhas“
2'-Hüdroksüatsetofenoon	Aldrich	99%
Antrakinoon	Prof. Tullio Ilometsa kollektsioonist	pole määratletud
9-Hüdroksüfluoreen	Aldrich	96%
Pürogallool	Sigma-Aldrich	ACS reagent
3-Metoksükatehhool	Alfa Aesar	98%
Süringool	Aldrich	99%
Sirelalaldehüüd	Aldrich	98%
Sirelhape	Sigma	≥95%
Gallushape	Harkovi keemiareagentide tehas	„puhas“
Vanilliinhape	Fluka	≥97%
4-Hüdroksübensoehape	Aldrich	99%
Trans-feruulhape	Aldrich	99%
4-Aniishape	Prof. Tullio Ilometsa kollektsioonist	pole määratletud
<i>n</i> -Heksaan	Sigma-Aldrich	≥97%
Tolueen	Sigma-Aldrich	99.9%
Klorobenseen	Fluka	99.5%
Diklorometaan	J.T.Baker	≥99.8%
Metüül- <i>tert</i> -butüüleeter	Sigma-Aldrich	99.8%
Isopropüülatsetaat	Sigma-Aldrich	98%
Bensüülalkohol	Sigma-Aldrich	99-100.5%
Tsükloheksanool	Sigma-Aldrich	99%
<i>n</i> -Butanool	Sigma-Aldrich	≥99.5%
<i>n</i> -Oktanool	Sigma-Aldrich	≥99%
Metanool	Sigma-Aldrich	≥99.9%
Ammoniumatsetaat	Fluka	≥99.0%

## **Infoleht**

### **Ligniini laguproduktide jaotuskoefitsientide eksperimentaalne määramine ja modelleerimine COSMO-RS meetodiga**

Käesoleva töö käigus uuriti eksperimentaalselt ligniini derivaatide ning lähedaste ühendite jaotuskoefitsiente 20 erinevas solvendipaaris ning hinnati saadud andmete abil COSMO-RS meetodi ennustusvõimet. Töö käigus leiti, et COSMO-RS meetodi võime ennustada jaotuskoefitsiente sõltub uuritavast aineist ning kasutatavast lahustist. Teisalt tuleb COSMO-RS meetodiga töötades olla ettevaatlik kasutades aineid, kus polaarsel asendusrühmal on mitu võimalust moodustada sisemolekulaarset vesiniksidet. Üldjuhul on eksperimentaalsete ja arvutuslike andmete kokkulangevus hea ja väärtused on omavahel tugevas korrelatiivses sõltuvuses.

Märksõnad: ligniin, jaotuskoefitsient, COSMO-RS

### **Experimental study and COSMO-RS modelling of partition coefficients of lignin decomposition products**

In this thesis partition/distribution coefficients of lignin derivatives and other similar compounds in 20 different solvent pairs were investigated experimentally. The obtained data were used to evaluate prediction capability of the COSMO-RS method. This thesis shows that the capability of COSMO-RS method to predict partition/distribution coefficients depends on the substance of interest and used solvents. On the other hand, when working with the COSMO-RS method one has to be careful if using substances where polar substituents have several possibilities of forming intermolecular hydrogen bonds. In general the agreement between experimental and computational data is good and values are strongly correlated.

Keywords: lignin, partition/distribution coefficients, COSMO-RS

CERCS kood: P300 analüütiline keemia

## **Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks**

Mina, Kertu Hernits,

(sünnikuupäev: 09.11.1993)

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose  
“Ligniini laguproduktide jaotuskoefitsientide eksperimentaalne määramine ja modelleerimine  
COSMO-RS meetodiga”,

mille juhendajad on Ivo Leito ja Sofja Tšepelevitš,

- 1.1.reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
- 1.2.üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu alates **01.01.2019** kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. olen teadlik, et nimetatud õigused jäävad alles ka autorile.
3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, **24.05.2017**